Lessons Learned from the Speech Recognition and Understanding project HAAWAII
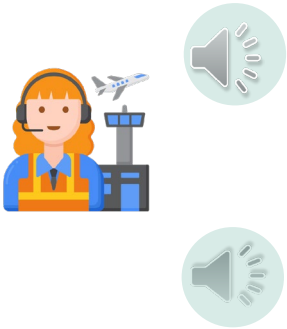
Mohsan Jameel, Hartmut Helmke

Institute of Flight Guidance, **German Aerospace Center (DLR)**
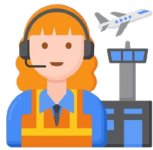
20th April 2023

The concept has been developed within the framework of SESAR 2020 and the Horizon 2020 research and innovation programme.

ATCo and Pilot Voice Communication

ATCo and Pilot Voice Communication

faroe line four five three good afternoon
radar contact squawk two seven six two

faroe line four five three after gunpa
cleared direct robur three eight zero

two seven sixty one squawk
faroe line four five three

gunpa robur flight level three eight zero
faroe line four five three

Readback Error

Readback Errors
Readback errors are seldom events and hard to detect

1-4%

Hearback Errors
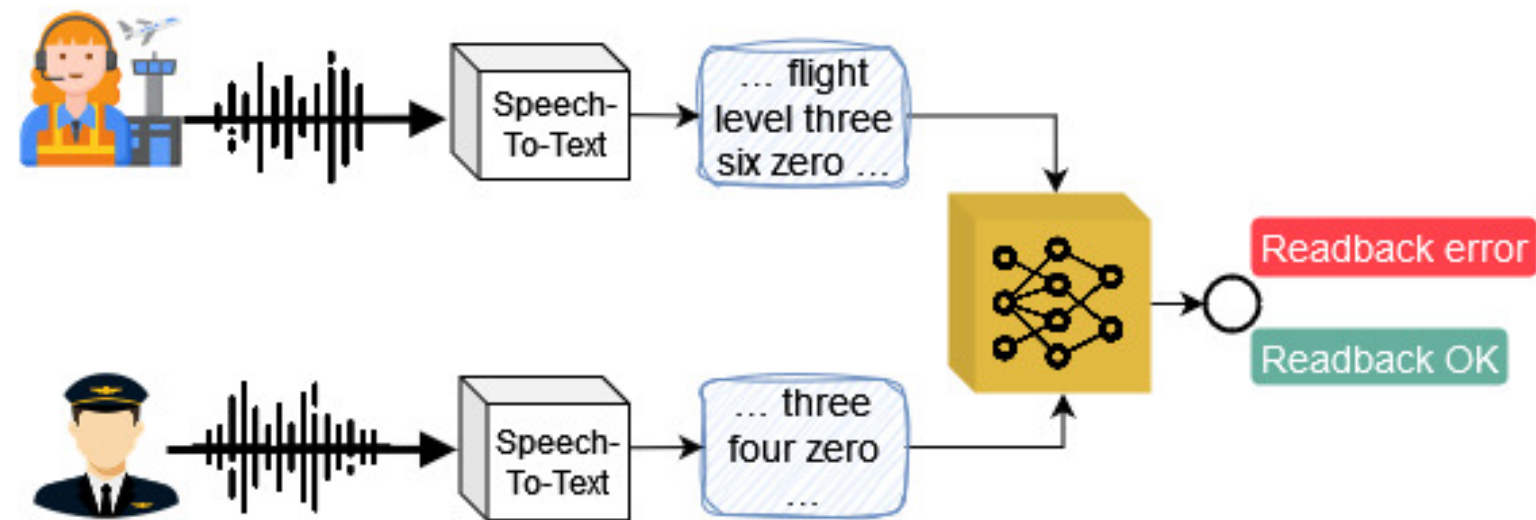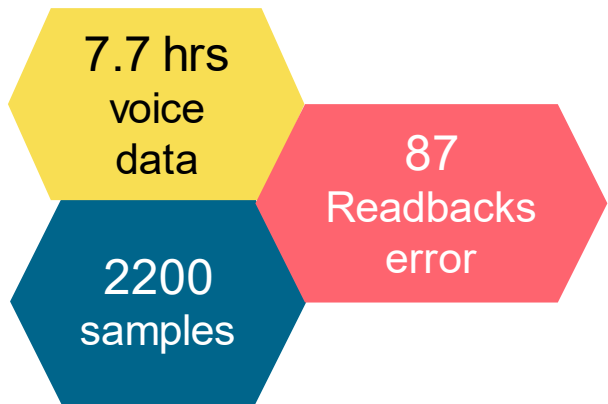Undetected readback errors are categorized as Hearback Errors

40-60%

Readback Error Detection Assistant (REDA)

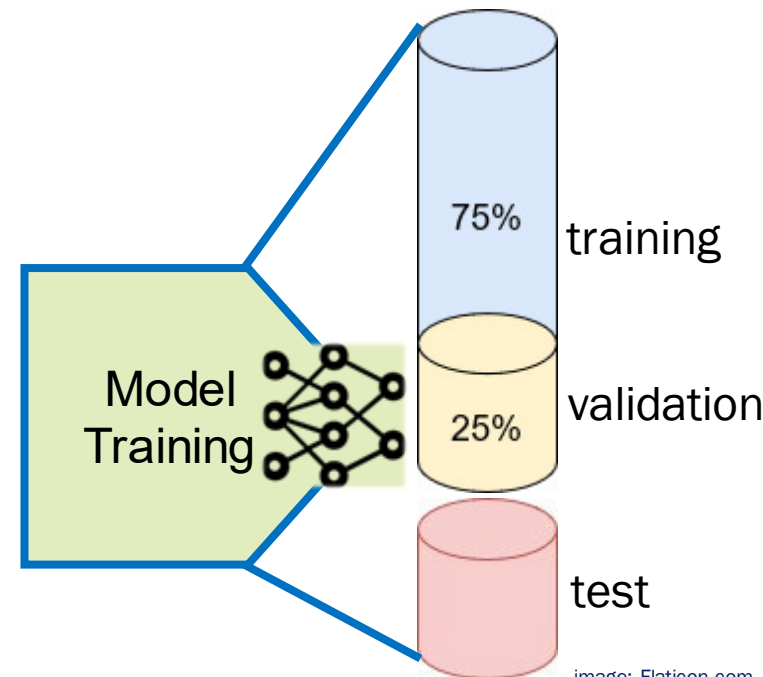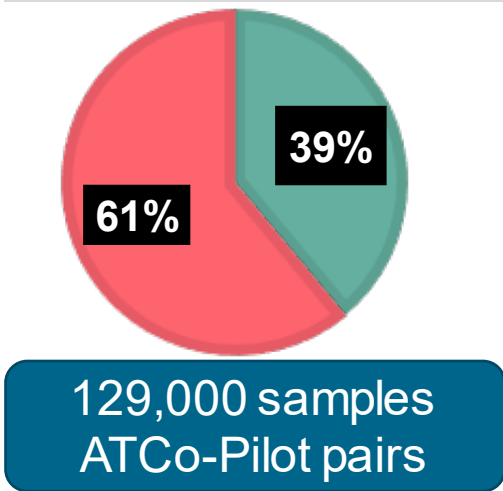# Data Driven Approach to Readback Error Detection Assistant

Speech-To-Text → … flight level three six zero …

Speech-To-Text → … three four zero …

Readback error

Readback OK

Voice data from NATS and Isavia

7.7 hrs voice data

2200 samples

87 Readbacks error

Generate Synthetic data

Synthetic training data

61%

39%

129,000 samples ATCo-Pilot pairs

Model Training

75% training

25% validation

test

image: Flaticon.com

DLR

Test Data: 87 readback errors / 2200 correct readbacks communications threads

	Absolute Number		Rates			
	Detected RB Errors	False Alarms	Detection	False Alarm	Accuracy	F1-Score
Data-Driven	37	133	**43%**	**78%**	92%	29%

Improve automation model ➡ Need Ops room data at large scale

Rule-Based Approach to Readback Error Detection Assistant
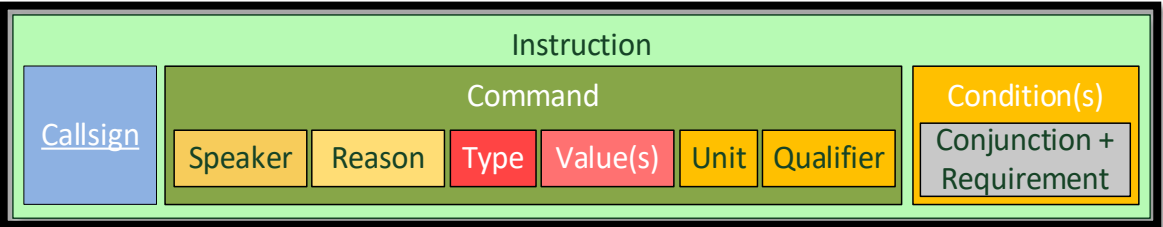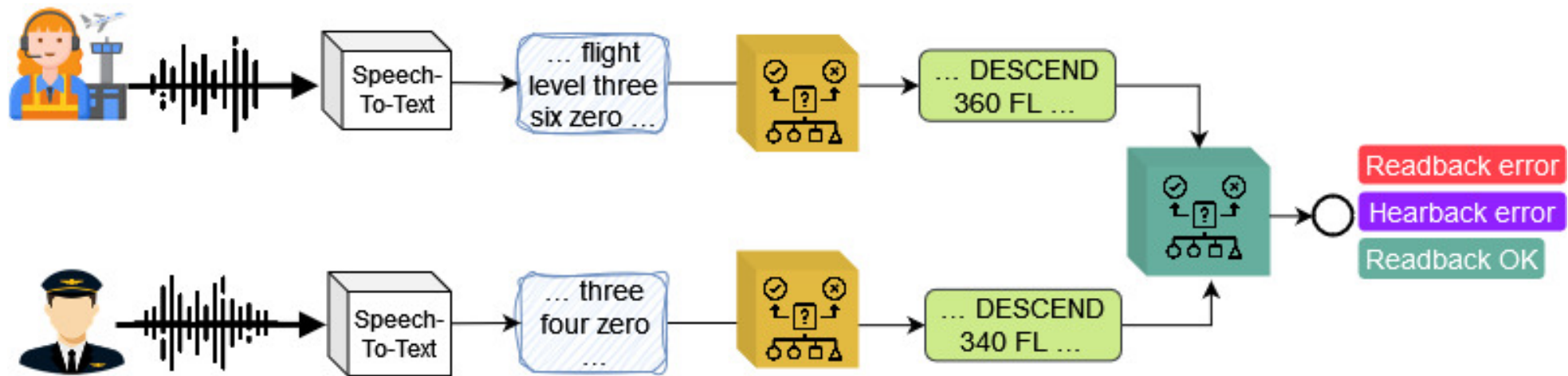
Ontology developed in SESAR PJ.16-04 by a consortium of European partners
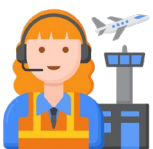
FLI453 GREETING
FLI453 INIT_RESPONSE
FLI453 SQUAWK 2762

faroe line four five three good afternoon
radar contact squawk two seven six two

FLI453 SQUAWK 2761

two seven sixty one squawk
faroe line four five three

image: Flaticon.com

Test Data: 87 readback errors / 2200 correct readbacks communications threads

	Absolute Number		Rates			
	Detected RB Errors	False Alarms	Detection	False Alarm ?	Accuracy	F1-Score
Data-Driven	37	133	43%	78%	92%	29%
Rule-Based	72	163	83%	69%	92%	45%
Combined	71	145	82%	67%	93%	47%

Improve automation model ➜ Need Ops room data at large scale

REDA: 82% detection rate > Human: 40-60% detection rate

Moving from Lab to Ops Room

- Decide on feedback about Readback errors
  - according to rules/manual ➔ phraseology is not "always" by the book
  - brought to the controller's attention ➔ decide on the context
  - communicated to the pilot ➔ decide on the type

- Show output at the supervisor position

- More data needed
  - Test data for offline evaluation
  - Feedback from the end user
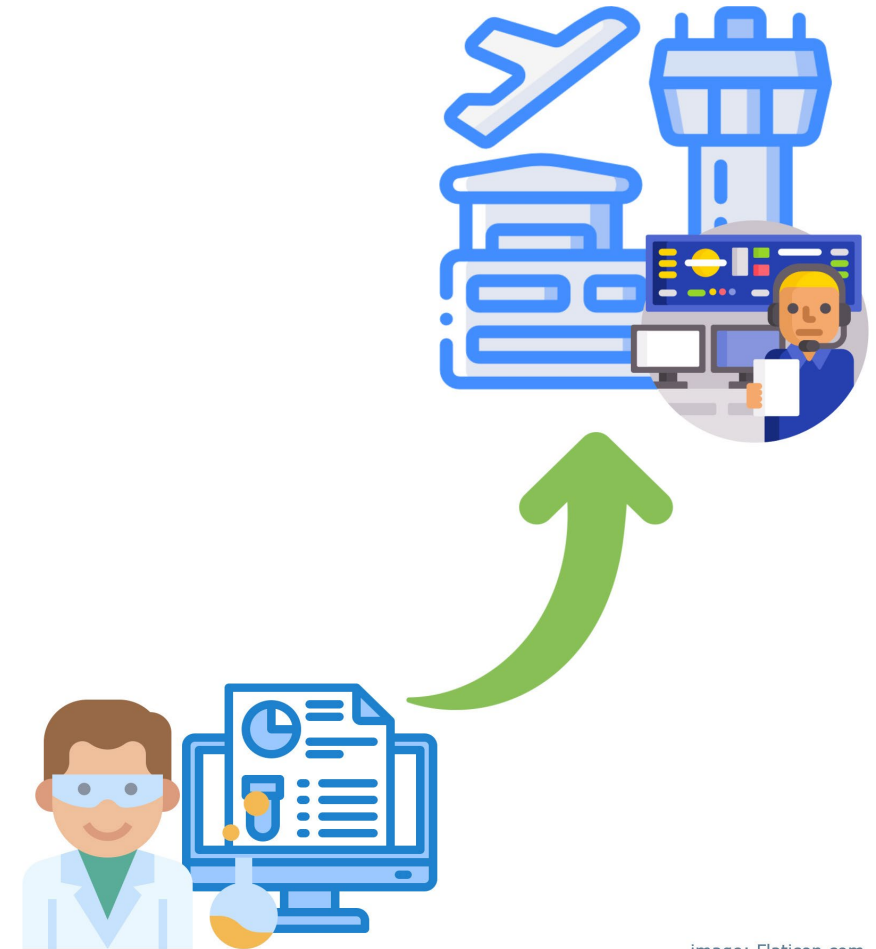 - Actual statistics of Readback errors (not just an approximation)

image: Flaticon.com

HAAWAII REDA prototype in Ops Room

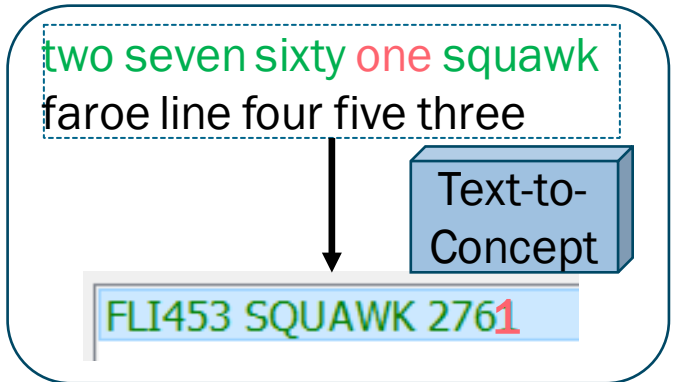**Date:** June, 29th 2022

Location: Reykjavik Ops Room

Event: REDA connected to **real-time** voice and surveillance data

Speech Recognition and Machine Learning Roadmap

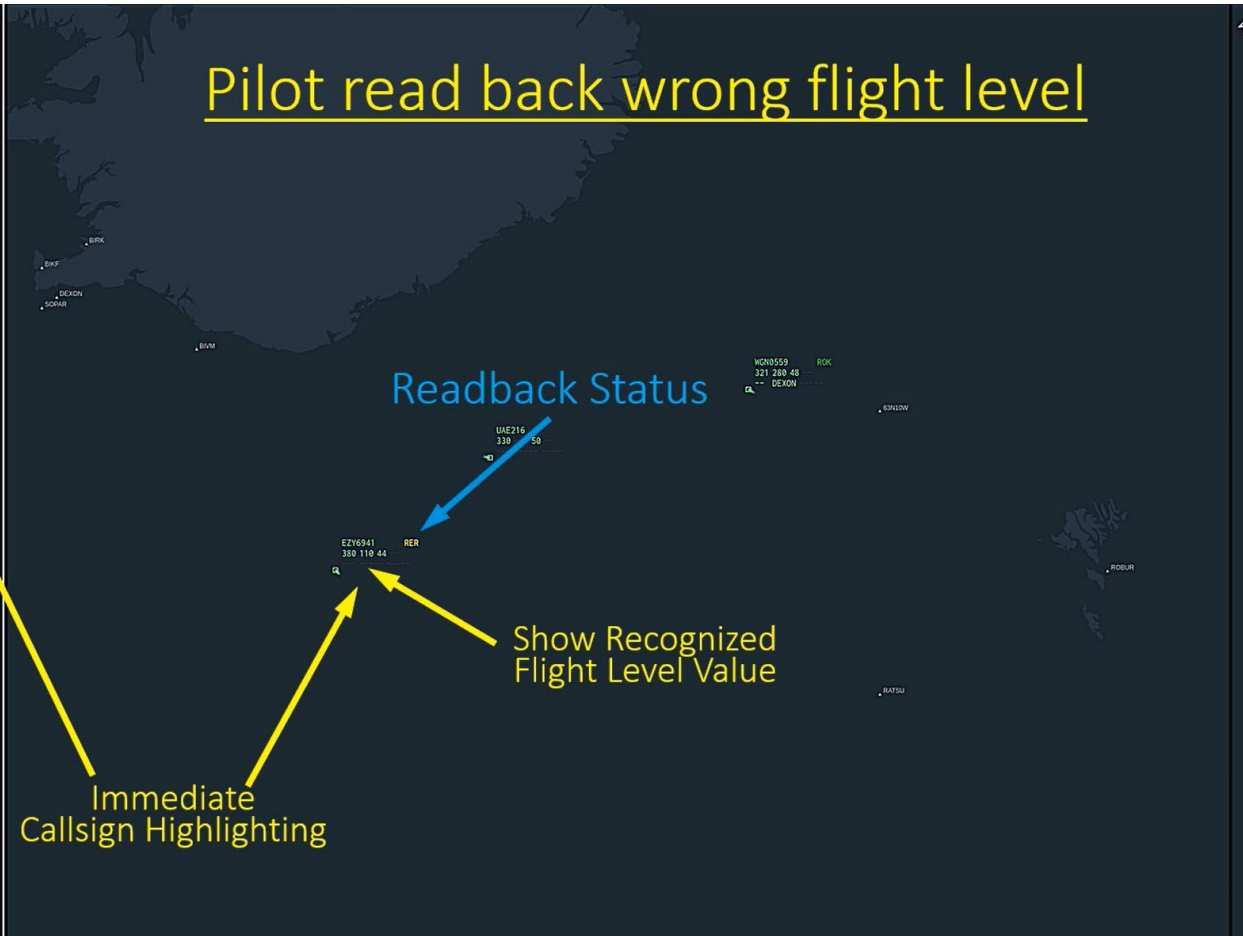If you are enthusiastic about supporting **ATCO's mission** through technology

Safety

Ecological | Efficient

ATCO's mission

Orderly | Expeditious

Economical

two seven sixty one squawk faroe line four five three

Text-to-Concept

FLI453 SQUAWK 2761

DLR offers to make source code available

HAAWAII

PJ.W2 PROSA

#97 Remote Tower

ATCO²

STARFiSH

AcListant

2015

MALORCA

2019

2022

Next ?

2014

AcListant ®-Strips

2018

CWP HMI

16-04

2020

2024

Pilot read back wrong flight level